

WHAT IS CLAIMED IS:

1. A speech recognition apparatus, comprising:

a feature extracting portion extracting a feature parameter by sliding, at least with different time width, a plurality of frames corresponding to time windows each having a prescribed length of time, over an input speech signal;

a storing portion storing standard pattern data in correspondence with phoneme patterns, respectively, of said input speech; and

a recognizing portion collating said feature parameter extracted by said feature extracting pattern with said standard pattern data to recognize a corresponding phoneme and to output a recognition result.

2. The speech recognition apparatus according to claim 1, wherein

said feature extracting portion successively increases time width for sliding said frame from a beginning part to an ending part of said input speech signal; and

said storing portion stores said standard pattern data corresponding to a pattern of time width with which said feature extracting portion slides said frame.

3. The speech recognition apparatus according to claim 1, wherein said feature extracting portion includes

a first fixed-frame-interval extraction processing portion extracting said feature parameter while sliding said frame with a first fixed time width, and

a second fixed-frame-interval extraction processing portion extracting said feature parameter while sliding said frame with a second fixed time width shorter than said first fixed time width; and

said standard pattern data include first standard pattern data corresponding to a first pattern of time width with which said first fixed-frame-interval extraction processing portion slides said frame, and a second standard pattern data corresponding to a second pattern of time width with

which said second fixed-frame-interval extraction processing portion slides said frame.

4. The speech recognition apparatus according to claim 1, wherein said feature extracting portion includes

a fixed-frame-interval extraction processing portion extracting said parameter while sliding said frame with a fixed time width, and

5 a variable-frame-interval extraction processing portion extracting said parameter while successively increasing time width for sliding said frame, from a beginning part to an ending part of said input speech; and said standard pattern data include

10 first standard pattern data corresponding to a first pattern of time width with which said fixed-frame-interval extraction processing portion slides said frame, and a second standard pattern data corresponding to a second pattern of time width with which said variable-frame-interval extraction processing portion slides said frame.

5. The speech recognition apparatus according to claim 1, wherein said feature extracting portion includes

a first fixed-frame-interval extraction processing portion extracting said feature parameter while sliding said frame with a first fixed time width, and

a second fixed-frame-interval extraction processing portion extracting said feature parameter while sliding said frame with a second fixed time width shorter than said first fixed time width; and

10 said standard pattern data include first standard pattern data corresponding to a first pattern of time width with which said first fixed-frame-interval extraction processing portion slides said frame, and a second standard pattern data corresponding to a second pattern of time width with which said second fixed-frame-interval extraction processing portion slides said frame;

15 said speech recognition apparatus further comprising an input selecting portion provided between said input speech signal

and said feature extracting portion, for switching destination of said input speech signal from said first fixed-frame-interval extraction processing portion to said second fixed-frame-interval extraction processing portion in accordance with a result of collation by said recognizing portion based on  
20 said feature parameter extracted by said first fixed-frame-interval extraction processing portion.

6. The speech recognition apparatus according to claim 5, wherein said first standard pattern data are related to time; said speech recognition apparatus further comprising  
an interpolating portion generating said second standard pattern  
5 data by interpolation, based on said first standard pattern data.

7. The speech recognition apparatus according to claim 6; wherein said first standard pattern data and said second standard pattern data are related to time; and  
each time point, at which said second fixed-frame-interval extraction  
5 processing portion slides said frame, corresponds to any of time points at which said first fixed-frame-interval extraction processing portion slides said frame.

8. The speech recognition apparatus according to claim 1, wherein said feature extracting portion includes  
a fixed-frame-interval extraction processing portion extracting said  
feature parameter while sliding said frame with a fixed time width, and  
5 a variable-frame-interval extraction processing portion extracting said parameter while successively increasing time width for sliding said frame, from a beginning part to an ending part of said input speech; and said standard pattern data include  
first standard pattern data corresponding to a first pattern of time  
10 width with which said fixed-frame-interval extraction processing portion slides said frame, and a second standard pattern data corresponding to a second pattern of time width with which said variable-frame-interval

extraction processing portion slides said frame;

15        said speech recognition apparatus further comprising  
an input selecting portion provided between said input speech signal  
and said feature extracting portion, for switching destination of said input  
speech signal from said fixed-frame-interval extraction processing portion  
to said variable-frame-interval extraction processing portion in accordance  
with a result of collation by said recognizing portion based on said feature  
20        parameter extracted by said first fixed-frame-interval extraction processing  
portion.

9. The speech recognition apparatus according to claim 8, wherein  
said first standard pattern data are related to time;  
said speech recognition apparatus further comprising  
an interpolating portion generating said second standard pattern  
5        data by interpolation, based on said first standard pattern data.